Leukemia (2005) 19, 1229–1238

www.nature.com/leu

Discrete serum protein signatures discriminate between human retrovirus-associated hematologic and neurologic disease

OJ Semmes[1,2], LH Cazares[1,2], MD Ward[1,2], L Qi[1,2], M Moody[1], E Maloney[3], J Morris[4], MW Trosset[5], M Hisada[3], S Gygi[6] and S Jacobson[7]

[1]Department of Microbiology and Molecular Cell Biology, Eastern Virginia Medical School, Norfolk, VA, USA; [2]Center for Biomedical Proteomics, Eastern Virginia Medical School, Norfolk, VA, USA; [3]Viral Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; [4]Metabolism Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; [5]Department of Mathematics College of William & Mary, Williamsburg, VA, USA; [6]Proteomics Center, Harvard Medical School, Boston, MA, USA; and [7]Viral Immunology Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA

The human T-cell leukemia virus type I (HTLV-I) is the causative agent for adult T-cell leukemia (ATL) and HTLV-I-associated myelopathy/tropical spastic paraparesis (HAM/TSP). Approximately 5% of infected individuals will develop either disease and currently there are no diagnostic tools for early detection or accurate assessment of disease state. We have employed high-throughput expression profiling of serum proteins using mass spectrometry to identify protein expression patterns that can discern between disease states of HTLV-I-infected individuals. Our study group consisted of 42 ATL, 50 HAM/TSP, and 38 normal controls. Spectral peaks corresponding to peptide ions were generated from MS-TOF data. We applied Classification and Regression Tree analysis to build a decision algorithm, which achieved 77% correct classification rate across the three groups. A second cohort of 10 ATL, 10 HAM and 10 control samples was used to validate this result. Linear discriminate analysis was performed to verify and visualize class separation. Affinity and sizing chromatography coupled with tandem mass spectrometry was used to identify three peaks specifically overexpressed in ATL: an 11.7 kDa fragment of alpha trypsin inhibitor, and two contiguous fragments (19.9 and 11.9 kDa) of haproglobin-2. To the best of our knowledge, this is the first application of protein profiling to distinguish between two disease states resulting from a single infectious agent.

Leukemia (2005) **19,** 1229–1238. doi:10.1038/sj.leu.2403781
Published online 12 May 2005
Keywords: myelopathy; proteomics; diagnostics

Introduction

Infection with the Human T-cell Leukemia virus type I (HTLV-I) can result in a number of disorders, including an aggressive T-cell malignancy, adult T-cell leukemia (ATL),[1] and a chronic, progressive neurologic disorder termed HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP). Globally, the estimated numbers of people infected with HTLV-1 approaches 20 million.[2,3] In endemic areas including southern Japan, the Caribbean basin and parts of western Africa and South America, where infection rates range from 2 to 30%, these diseases are the major causes of mortality and morbidity.[4] For example, ATL is the most common form of non-Hodgkin's lymphoma in Southwestern Japan and HAM/TSP is the most common form of neurodegenerative disorder in central Brazil.[2–4] Although there are adequate methods for determining if people are infected with HTLV-I, there are no diagnostic tools available for predicting disease status. The existing 'diagnostic markers' for ATL currently are immunoassay for *HTLV-I* gene products, HTLV-I-specific antibody production, and detection of HTLV-I DNA.[5] These parameters have little discriminating value among HTLV-I diseases, clinical subtypes, or severity of disease. There has been some utility of immune activation markers as predictors of outcome, which, although effective within ATL, are not diagnostic for HAM/TSP or other HTLV-I-related diseases as a whole.[6]

Survival of patients with ATL is grim and has a mean life expectancy of 9 months from the time of diagnosis.[7] The treatment of patients with ATL has been largely unsuccessful, resulting in little consensus on the best treatment regimen. The situation is no better for HAM/TSP in that therapies have been confined to approaches shown to be successful with other neurodegenerative diseases such as multiple sclerosis.[8] Although several new strategies for immune suppression are being initiated, these treatments will not eliminate the infected T-cell population and would presumably require long-term immunoregulation for this chronically progressive disorder.

In light of the lack of effective treatment in later disease states, it would be useful to know early in infection who may be at high risk for developing disease. In the early stages, aggressive treatments to eradicate the virus may prove effective since viral load is believed to be low. These treatments, however, may have serious side effects that would contradict their use in the HTLV-I-infected asymptomatic population at large. Additionally, conversion from HAM to ATL, although very rare, is a complication that is difficult to establish short of lymphadenectomy. The need for clear distinction between ATL and HAM via noninvasive approaches to identify the highest risk group would have obvious merit.

Surface Enhanced Laser Desorption Ionization (SELDI) time-of-flight mass spectrometry (TOF-MS) has received tremendous attention in the field of cancer diagnostics.[9–11] The primary advancement that sets SELDI apart from other instrumentation utilized for protein expression profiling is the employment of the ProteinChip, an ionization source surface that has been chemically treated to incorporate an affinity layer, which is otherwise similar to Matrix-Assisted Laser Desorption Ionization (MALDI). Previously, we established the utility of SELDI-TOF-MS in disease diagnosis by application to transitional cell carcinoma in bladder.[12] Subsequently, others and we have shown that the application of learning algorithms to the time-of-flight data results in highly accurate diagnostic approaches for

Correspondence: Dr OJ Semmes, Department of Microbiology and Molecular Cell Biology, Eastern Virginia Medical School, 700 West Olney Road, Center for Biomedical Proteomics, Norfolk, VA 23507, USA. Fax: +1 757 446 5766; E-mail: semmesoj@evms.edu
These studies were supported by a Translational Research Grant from the Leukemia and Lymphoma Society to OJS
Received 10 January 2005; accepted 21 March 2005; Published online 12 May 2005

prostate, breast, and ovarian cancer.[9,10,13,14] One of the challenges of successful application of serum protein expression profiling to early cancer detection is the ability to detect potentially rare protein events that signal early tumor formation. In this study, we have applied this technology to a human infectious disease that results in both hematological and neurological disorders, each of which might be expected to result in significant protein changes in blood. SELDI-TOF-MS coupled with a regression tree analysis was used to distinguish between HTLV-I-uninfected healthy individuals and HTLV-I-associated ATL or HAM/TSP.

Methods

Patient population and samples

The patient population consists of 42 individuals with ATL, 50 patients with HAM/TSP, and 38 healthy uninfected individuals seen at the National Institutes of Health, Bethesda, MD, on a number of IRB-approved studies. The ATL patients had a mean age of 44.3 years (range 25–69) and a male/female ratio of 0.73. The ethnicity of the cohort was 60% Afro-Caribbean, 17% African-American, 10% African, 10% Caucasian, and 3% Hispanic. The ATL patients were diagnosed as chronic (10%), acute (50%), smoldering (10%), and clinical indistinct (30%). In all, 50% of the patients were previously untreated, 33% received prior chemotherapy or combined AZT/interferon-alpha, and the treatment status was unknown to the remaining patients. The diagnosis and classification of HTLV-I-associated ATL was made using the WHO classification and Shimoyama criteria.

The HAM/TSP patients had a mean age of 49.7 years (range 13–80), a male/female ratio of 0.45, and a mean duration of disease of 9.3 years (range 2–23). The healthy control group had a mean age of 40.5 years (range 19–56), with a male/female ratio of 0.32. The diagnosis of HAM/TSP was assessed according to the WHO guidelines. All HAM/TSP patients were evaluated at the NIH Clinical Center for progressive spastic paraparesis and were serologically shown to be HTLV-I Western blot reactive. The range of disability as assessed by the expanded disability status scale (EDSS) was 2.0–8.0, with a median of 5.5. Serum samples were obtained prior to any immunomodulatory therapy, aliquoted and frozen for future proteomic studies.

Sample acquisition and preparation

Whole blood was drawn from individuals following proper consenting under NIH protocol 98N-047. The blood was collected in a 10 cm^3 serum separator vacutainer tube and centrifuged for 15 min at 1500 *g* to separate out the serum fraction. Serum was immediately transferred to ice. The samples were then aliquoted into 500 μl fractions and stored at -70 to -80°C immediately following phlebotomy. Each fraction was limited to a single freeze–thaw prior to analysis. All serum samples were collected following the same protocols as described above.

SELDI-TOF-MS

We employed an in-house program to assign samples in a randomized matrix pattern to prevent bias between triplicate or clinical status and chip spot position. All samples were processed in triplicate and the arrayed chips were read in a 48-h period. The matrix codes were assigned by an individual separate from the team that processed the samples so that each phase of the study was blinded with respect to the operator. The code was broken during the classification stage.

All sample and chip processing was performed using the BioMek 2000 workstation (Beckman) adapted to run the ProteinChip. Briefly, 20 μl of serum is pretreated with 8 M urea, 1% CHAPS, and placed on a MicroMix Shaker (DPC) for 10 min at 4°C. A further dilution is made in 1 M urea, 0.125% CHAPS, and PBS. A volume of the diluted serum (100 μl) is then added to the ProteinChip® with the aid of a bio-processor. ProteinChips® are then incubated at room temperature for 30 min with shaking, followed by washes of PBS and water. Arrays were allowed to air dry and then spotted with sinapinic acid in 50% (v/v) acetonitrile and 0.5% (v/v) trifluoroacetic acid. The ProteinChips® were analyzed using the SELDI ProteinChip® System (PBS-II, Ciphergen Biosystems, Inc.). Spectra were collected by the accumulation of 192 shots in the positive mode. The protein masses were calibrated externally using purified peptide standards (Ciphergen Biosystems, Inc.) Instrument settings were optimized using a pooled serum standard.

Data analysis

Before analysis, the data were divided into two sets as follows: a training set, which consisted of 32 ATL, 40 HAM, and 28 normals, and a test set consisting of the remaining samples. Spectra were analyzed with the Ciphergen ProteinChip® software (version 3.0) and normalized using total ion current. Peak labeling and clustering were performed using the Biomarker Wizard tool in the software, exported into a spreadsheet, and the intensity values for each peak were averaged for triplicate samples. These processed spectral data were then analyzed by Classification and Regression Tree (CART) using the BioMarker Patterns Program (Ciphergen Biosystems) to develop a classification tree.

In the case of analysis of the second validating test data set, 10 ATL, 10 HAM, and 10 controls were run in triplicate as described above. However, the peak identification step varied from that employed with the first data set. Specifically, the peak list generated from the first study was used to 'cluster' or identify peaks from the validating test set using a 0.2% mass window bin. The averaged normalized intensity values for each of the 'peaks' in the validating test set data that matched the 'peaks' in the decision algorithm were applied to the 'cut-off' values of the decision node.

Classification algorithm

Details regarding CART and its application to the analysis of protein expression profiling data have been described elsewhere.[11,15,16] Classification trees were constructed using the training set, and a V-fold cross validation process. Multiple classification trees were generated using this process, and the best performing tree was chosen for further testing using the cross-validation results. The reported accuracy of the selected classification tree was determined by challenging with the blinded test set.

Protein isolation and partial purification

To purify peaks observed in the SELDI spectra, 200 μl of serum samples with and without the peaks of interest were diluted with the same volume of 8 and 1 M urea/CHAPS as above. The total

volume of the scaled-up sample was 7.5 ml. Samples were then added to Talon-spin immobilized metal affinity columns (Clontech, Palo Alto, CA, USA), which were charged with $CuSO_4$ to mimic the IMAC Cu^{2+} surface used on the SELDI chips. After washing with PBS, the bound proteins were eluted from the column as recommended according to the manufacturer's protocol, with an elution buffer containing imadiazole. The total elution volume was 300 μl. A parallel sample was eluted in high pH and subsequently neutralized prior to analysis by SELDI-TOF-MS. A volume of 12 μl of each of the eluates representing approximately 24 μg was run on a NuPAGE 4–12% Bis–Tris gel under reducing conditions. Gels were stained with the SilverQuest staining kit (Clontech).

Alpha-1-antitrypsin immunodepletion

In all, 20 μl of serum was processed in urea/CHAPS and diluted in PBS using the same volumes as those employed for the SELDI analysis above. The diluted serum was incubated with 100 μl protein A-Sepharose beads (BSA treated and washed in urea buffer) for 2 h at 4°C. The supernatant was collected after a brief spin and the sepharose beads containing bound serum IgG were saved. A volume of 6 μl (1 mg/ml) of rabbit anti-human alpha-1-antitrypsin antibody (DakoCytomation Carpinteria, CA, USA) was then added to the 'IgG-cleared' serum supernatant. PBS was added to bring the final volume of each sample to 1 ml and samples were then incubated overnight at 4°C. Protein A-Sepharose beads (150 μl) were then added and the samples were incubated at 4°C for 1–2 h. Following centrifugation, the sepharose beads containing the bound Ag–Ab complex and the 'Ag-cleared' supernatant were collected. The Ab–Ag complex-bound sepharose beads were washed four times with urea buffer in bulk. Elution of the beads was performed by incubation with 100 μl of elution buffer (0.15 M NaCl, 0.1 M acetic acid, pH 3.0) and shaking for 30 min at room temperature. Samples were spun briefly to collect the elution and 50 μl Tris, pH 9.0, was added to neutralize. SELDI analysis was performed using IMAC3-Cu^{2+} chips, with 100 μl of the fractions placed on each array for 30–60 min.

In-gel trypsin digest and LC-MS/MS analysis

SDS-PAGE gel slices were cut into 1–2 mm cubes, washed 3 × with 500 μl ultra-pure H_2O, and incubated in 100% acetonitrile for 45 min. If the gel was silver stained, the stain was first removed with SilverQuest™ destaining solution following the manufacturer's instructions. The material was completely dried in a speed-vac and rehydrated in a 12.5 ng/μl modified sequencing grade trypsin solution (Promega) and incubated in an ice bath for approximately 45 min. The excess trypsin solution was then removed and replaced with enough 50 mM ammonium bicarbonate, pH 8.0, to cover the gel slice, typically 50 μl. The digest was allowed to proceed overnight at 37°C. Peptides were extracted 2 × with 25 μl 50% acetonitrile, 5% formic acid, and dried in a speed-vac. The peptides were resuspended in 5% acetonitrile, 0.5% formic acid, 0.005% heptafluorobutyric acid (Buffer A), and 3–6 μl applied to a 70 μM ID, 15 cm Magic C18 reverse-phase capillary column. Peptides were eluted with a 5–80% acetonitrile gradient (Buffer A + 95% acetonitrile) and analyzed on a ThermoFinnigan LCQ DECA XP Ion Trap tandem mass spectrometer in positive ion mode. For each scan, the three highest intensity ions were subjected to MS/MS analysis. Sequence analysis was performed with Sequest™ using an indexed human subset database of the nonredundant protein database from NCBI.

Results

Reproducibility of SELDI-TOF profiling of cohort samples

A key aspect of any clinical approach for reliable disease diagnosis and early detection is reproducibility. We have

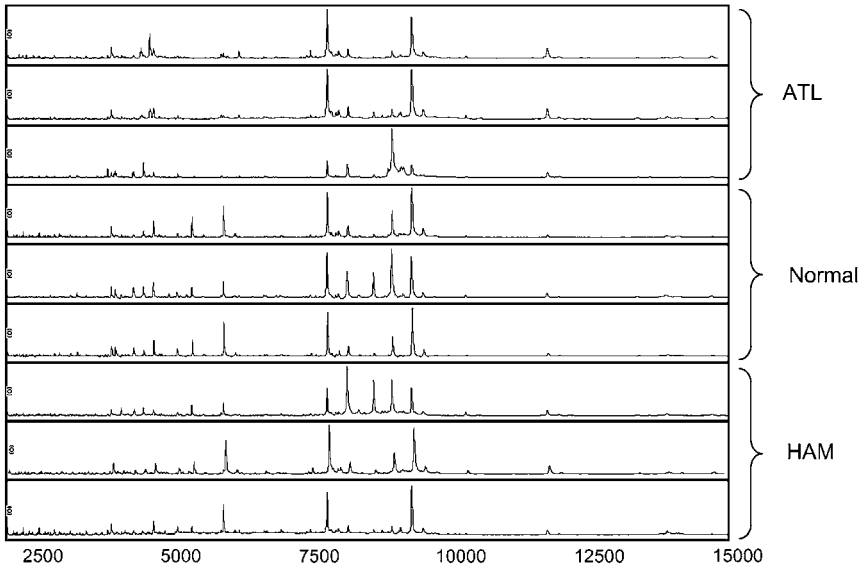

Figure 1 Reproducibility of SELDI profile. (a) The same serum sample was run 6 months apart. The top panel (1) is the SELDI spectra initially obtained, and the bottom panel (2) is the spectra obtained when the sample was run 6 months later on the same instrument. (b) Three separate samples from each class were assayed and the resulting spectra were normalized to total ion current. The expression values are displayed as relative intensity on the same absolute scale. The classes are ATL (one), HTLV-associated myelopathy (22), and normal control (Normal).

established optimal performance parameters beyond the standard calibration steps that enable us to optimize and monitor the performance of the instrument.[17] Using these parameters, we have successfully synchronized multiple instruments and have maintained this synchrony over a 3-year period.

In Figure 1, we show spectra from three separate representative individuals from each class. The variation between identical sample spectra is less than 0.1% for mass designation and expression amplitude displayed a CV of 15–20%.

Identification of differentially expressed m/z *peak values*

After calibration and normalization of the entire data set, consistent peak sets or clusters present in at least 10% of the samples were generated based on a mass window of $\pm 0.2\%$ of mass. Intensity values were averaged between the triplicates and reported for each peak set and differences between groups identified. Thus, peaks were identified based on being greater or lesser expressed in ATL, HAM, or healthy individuals. Using this selection process, a number of potential classifier peaks were found for ATL and HAM/TSP, respectively. Table 1 is a partial list of peaks that have been demonstrated to have discriminatory power, and are ranked by *P*-value. We show some examples of peaks that represent the differential expression observed between groups. Some of the peaks were present in normal and HAM/TSP sera, but not in ATL (for example, a peak at 3977 *m/z* shown in Figure 2a and a peak at 5360 *m/z* shown in Figure 2b), while others were only detectable in ATL sera (for example, peak 4597 *m/z* shown in Figure 2a). A peak at 11 738 *m/z*, which was present in most samples, was overexpressed in ATL sera compared to either HAM/TSP or control (Figure 2b).

Development of a decision tree classification algorithm

The verification of the utility of individual peaks as diagnostic biomarkers was addressed using CART analysis. The CART software uses the ranked peaks and evaluates for the ability to distinguish between classes and then applies fit-value assignments to each class. We directed the algorithm to segregate via three comparison schemes: ATL *vs* N; ATL *vs* HAM + N; and HAM *vs* N. The tree development is based on establishing simple cutoffs for expression values of selected peaks. Each of the decision nodes partitions the data into two groups, which are subsequently divided at secondary nodes until the optimal tree is 'grown'. The formed trees are subjected to 'pruning'; basically branches are removed and the 'cost' of the removal determined to establish a minimal tree size. The optimal tree that does not over-fit the data is then chosen using a cross-validation approach.

We have illustrated this process by presenting the actual relative values in a scatter plot for each decision peak in the ATL *vs* N tree (Figure 3). In this decision tree, peaks at 6136 and 11 768 *m/z* were able to distinguish between ATL and N effectively. However, the best separation was achieved with the combined use of these peaks. The ability to distinguish ATL from N was achieved with 97% sensitivity and 96% specificity, using a V-fold cross validation of the training set. The blinded test set correctly classified 90% (9/10) of ATL and 80% (8/10) of N.

Although it is useful to distinguish ATL from healthy individuals, the most useful clinical separation is between ATL and non-ATL (both HAM and N). In order to achieve this separation, we first addressed two didactic trees, ATL *vs* HAM + N and HAM *vs* N. The application of the regression tree analysis resulted in the trees shown in Figure 4. Training with cross-validation for the ATL *vs* HAM + N resulted in 91% sensitivity and 87% specificity. The blinded test set achieved 80% correct classification of ATL (8/10) and 85% correct classification of HAM/TSP and N (17/20). The subsequent decision tree training with cross-validation, for

Table 1 Most differentially expressed peaks

Cluster	*P-value*	*m/z average*
P-values (ATL vs Normal)		
10	0.0000000002	4666.39[a]
13	0.0000000012	5360.86[a]
15	0.0000000019	5930.34[a]
3	0.0000000108	3904.93[a]
16	0.0000000152	6136.67[a]
35	0.0000002772	9323.90[a]
36	0.0000020816	9531.79[a]
22	0.0000037021	7796.78[a]
25	0.0000078129	8004.92[a]
41	0.0000085632	11 769.35[b]
59	0.0000140854	78 926.34
42	0.0000261106	11 950.48[b]
P-values (HAM vs Normal)		
3	0.0000016437	3904.93[a]
59	0.0000027555	78 926.34[a]
57	0.0000137336	66 714.96[a]
51	0.0000203647	44 472.08[a]
50	0.0000362733	40 026.73[a]
58	0.0000490222	74 820.53[a]
45	0.0000659447	14 122.64[a]
33	0.0000915497	9174.99[b]
31	0.0000949140	8966.85[b]
56	0.0002061587	60 989.01[a]
49	0.0002531877	33 542.23[a]
48	0.0004617588	28 159.18[a]
P-values (ATL vs HAM[c])		
10	0.0000143257	4666.39[a]
13	0.0000313181	5360.86[a]
16	0.0000313181	6136.67[a]
15	0.0000507364	5930.34[a]
42	0.0001128584	11 950.48[b]
41	0.0005834305	11 769.35[b]
35	0.0050737586	9323.90[a]
22	0.0094501735	7796.78[a]
46	0.0094501735	14 737.49[b]
18	0.0121501941	6880.98[b]
43	0.0132898026	13 396.08[b]
40	0.0142055150	11 131.00[a]
P-values (ATL vs N+HAM)		
10	0.0000000002	4666.37[a]
13	0.0000000032	5360.91[a]
15	0.0000000059	5930.42[a]
16	0.0000000089	6136.71[a]
42	0.0000024228	11 950.49[b]
41	0.0000044566	11 769.44[b]
3	0.0000056628	3904.94[a]
35	0.0000063774	9323.96[a]
22	0.0000137628	7796.86[a]
25	0.0000368387	8004.93[a]
36	0.0000779729	9531.85[a]
20	0.0004684201	7595.06[a]

[a]Underexpressed in disease state.
[b]Overexpressed in disease state.
[c]In the case of ATL *vs* HAM, ATL is the disease state.

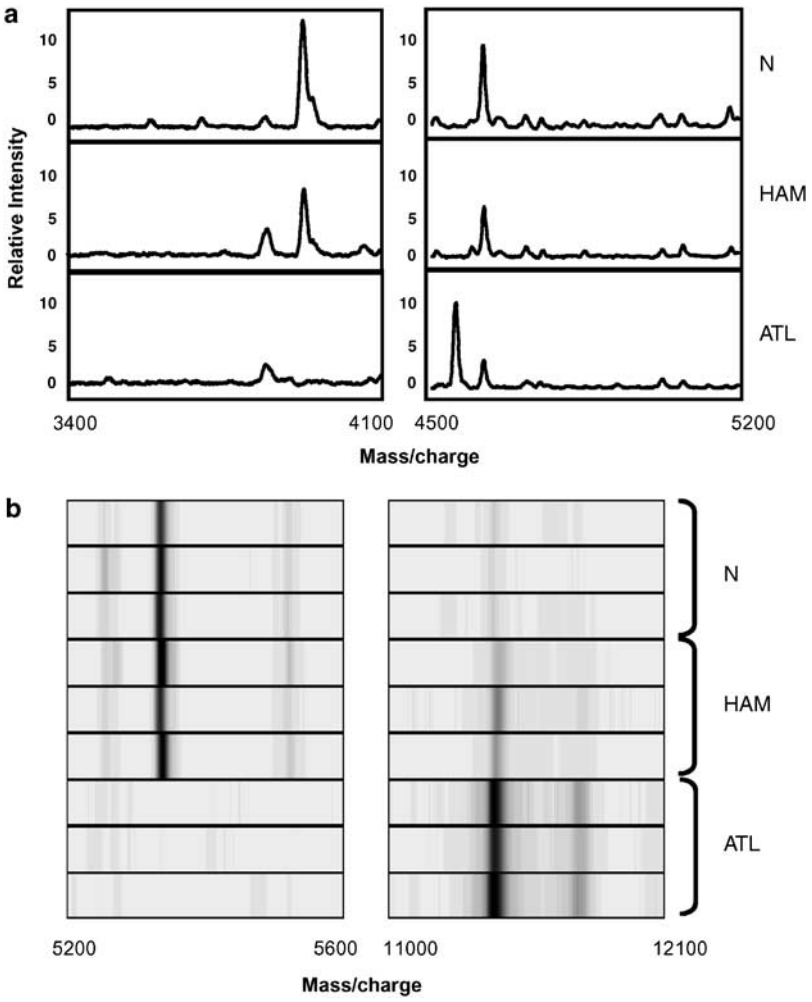

Figure 2 SELDI data showing class-specific peaks. (a) A peak absent in ATL at *m/z* 3907 is shown in the left panel and a peak unique to ATL right panel is visible at *m/z* 4597. (b) Gel-view representation of SELDI spectra showing a peak absent in ATL (*m/z* 5360) and a peak expressed at different levels in all classes (*m/z* 11 738).

HAM *vs* N, resulted in a sensitivity of 85% and specificity of 93%. The results of the blinded test set in this group achieved 90% correct classification of HAM and 90% correct classification of N. When we combined the two decision trees in series and challenged them with all the blinded test samples, we achieved 80% correct classification of ATL (8/10), 70% correct classification of HAM (7/10), and 80% correct classification of N (8/10). The decision structure for the combined tress is shown in Figure 4c. These results were achieved using a simple CART approach, and we anticipate that improvements in the node decisions (multiple peaks) as well as boosting approaches such as we have successfully applied to other SELDI data sets[14] would dramatically improve the correct classification rates. However, the simplicity of this design suggests that the protein peak profiles are significant in discerning HTLV-I disease.

Evaluation of the expression of 11 768 m/z *peak across populations and classification of a separate blinded validation test group*

To demonstrate the ability of the 11 768 and 6136 *m/z* peaks to segregate patients diagnosed with ATL from healthy individuals, we used the absolute cutoff values for these peaks in a scatter plot representation. In Figure 5a, the contribution of each of these peak values to the discrimination between groups can be visualized. The single value of the 11 768 *m/z* peak alone gave good separation between ATL and healthy individuals, reflecting the fact that it is overexpressed in ATL. We also tested the robustness of the classifier across patient samples. In this exercise, we tested the established combined decision tree described earlier (Figure 4c) with data from a follow-up study run 1 year later. In this study, the recruitment criteria and case groups were the same, but included serum from an entirely different group of patients. This completely separate blinded validation test group consisted of 10 ATL, 10 HAM, and 10 healthy individuals. When the data were applied to the decision structure, we achieved a correct classification rate of 73%: with 80% of ATL, 89% healthy individuals, and 50% of HAM/TSP correctly classified. Interestingly, the overall effectiveness of the decision algorithm with respect to class was retained with HAM/TSP being misclassified as healthy individuals. In addition, when we examined the 11 768 *m/z* peak in the test data set, the ability to discriminate using the same cutoff value was retained (Figure 5b). Thus, although validation upon a larger population

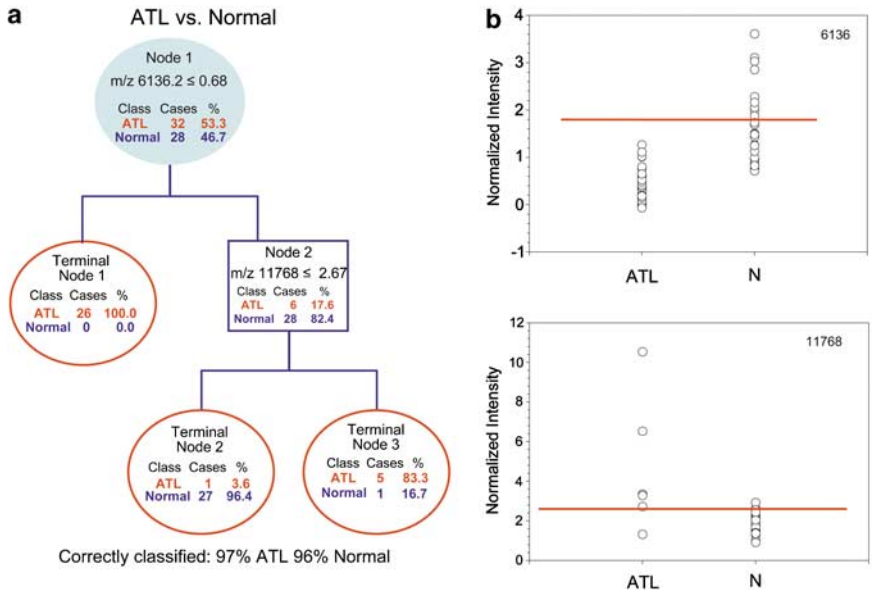

Figure 3 Application of CART and development of a decision tree for distinguishing ATL from normal healthy individuals (N). (a) The decision tree graph. The primary decision nodes are *m/z* 6136 and *m/z* 11 768. The terminal nodes are indicated as red circles. The training with v-fold cross-validation resulted in 97% sensitivity and 96% specificity. (b) Scatter plot diagram for the primary and secondary tree decision nodes. A scatter plot depiction of the relative variation in expression values for each decision event for the training set is shown. The decision cutoff is represented by a horizontal line; samples are referred to the secondary node, which is applied based on the value displayed.

is needed, the chosen peak expression pattern appears to be robust between study populations. The slight reduction in overall correct classification of all the three disease groups was due to the reduced segregation between the HAM/TSP and control groups and should be improved when this study is reproduced in the larger population.

Sequence identification of ATL 'diagnostic' peaks

As the application of protein expression profiling such as we describe here becomes more widespread, determining the identity of the protein/peptides that are differentially expressed between individuals with a disease and healthy control will allow for development of immunodiagnostic approaches as stand-alone or complimentary to profiling. In addition, characterizing these peaks will help validate the approach by identification of the specific proteins and molecular events involved in disease onset and progression. We have developed a purification scheme for identifying the SELDI-designated protein peaks. Two samples, one overexpressing and one underexpressing the desired peak, were identified for isolation and purification by examining the SELDI profiles of all samples (Figure 6). The isolation and identification strategies were applied to the pair so that a didactic comparison is available throughout the purification scheme. Specifically in this case, the paired samples were first reacted with a Cu^{2+} affinity column that emulates the on-chip affinity process of the SELDI-TOF-MS IMAC ProteinChip. This step also greatly reduces serum globulins without application of secondary fractionation steps. The affinity eluate was confirmed to contain the target peaks via reapplication onto SELDI-TOF-MS and then applied to a single dimension SDS-PAGE and silver stained. The visible differentially expressed bands within the targeted size range were excised in pairs and analyzed by capillary liquid chromatography coupled to electrospray tandem mass spectrometry.

Using this approach, we identified the diagnostic peaks as fragments of common serum proteins. Specifically, we identified the 11.7 *m/z* SELDI peak that was overexpressed in ATL as being a discrete fragment of alpha-1-antitrypsin (Figure 7a). We also identified two other peaks overexpressed in ATL, 11 950 and 19 872 *m/z*, as being two contiguous fragments of haptoglobin-2. Each of these peptide identities was supported by sequence coverage consistent with the proposed mass. Subsequent verification that the identified protein/peptide gives rise to the observed 'diagnostic' peak was accomplished by immunodepletion and is shown for alpha-1-antitrypsin (Figure 7b).

Discussion

A particular shortcoming of many studies involving the discovery of diagnostic biomarkers is the lack of preliminary evidence suggestive of the 'robustness' of the marker discrimination. In this study, we have structured our analysis to incorporate separate training and test data sets. This practice provides some insight into the expected true error of the classification algorithm. In addition, we have provided a second 'validating' test set derived from a separate patient cohort with similar disease classes. The ability to retain an equivalent accuracy in discrimination across the two patient cohorts is a measure of the robustness of the chosen biomarker. Thus, although these findings need to be verified in larger cohorts of patients, the results suggest that the profile generated from the selected proteins/peptides is class-specific.

The ability to identify a serum signature profile in a human infectious disease, to our knowledge, has not been reported previously. Moreover, this serum profile was able to distinguish a clinical outcome of a human retroviral infection. While serum signature profiles have been demonstrated for solid tumors (prostate, breast, ovarian cancers), no serum protein profiles

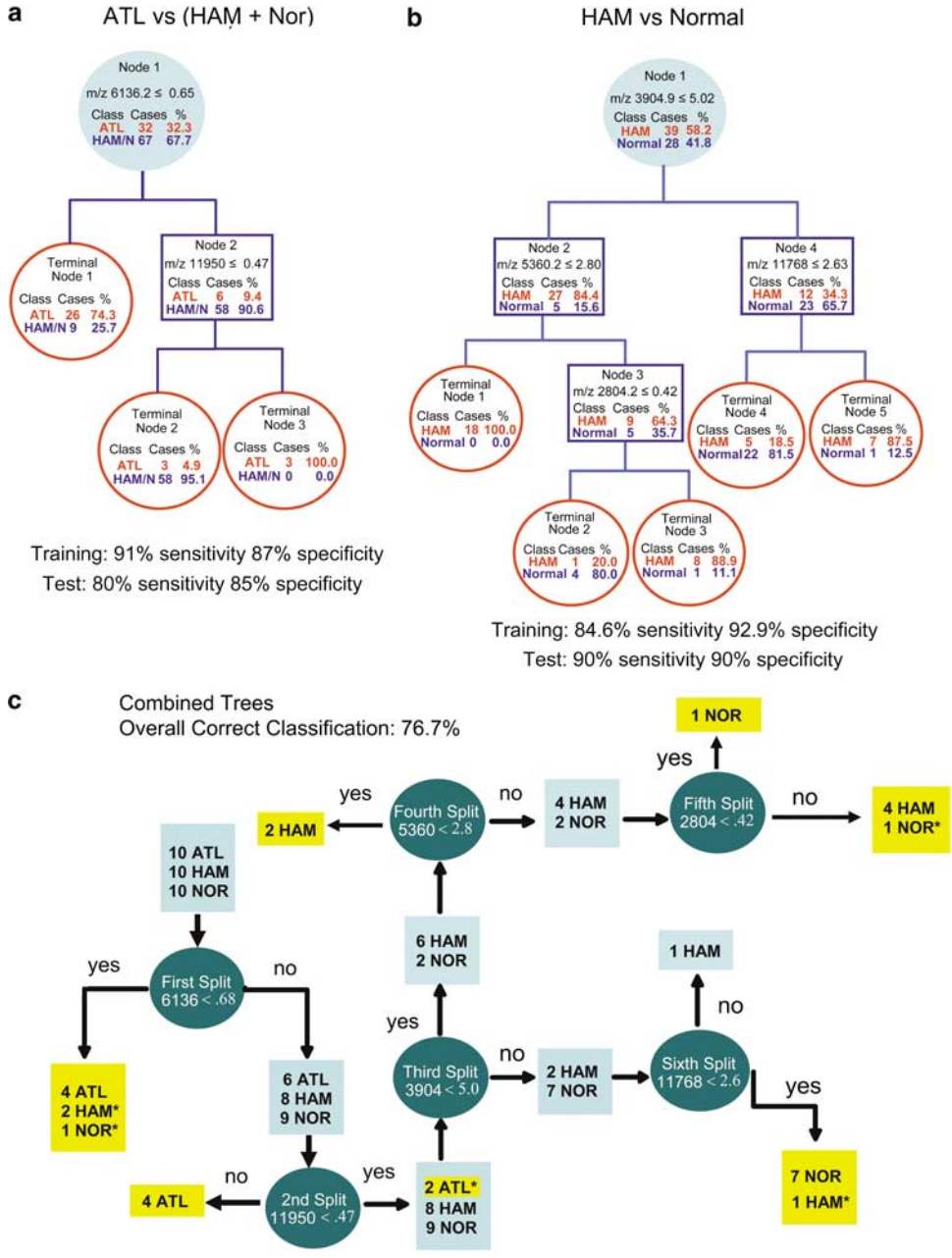

Figure 4 Combined decision tree for separation of ATL, HAM, and normal healthy individuals. (a) Decision tree developed to separate ATL from the combined class of HAM and healthy individuals. The *m/z* values and the intensity cutoff values for each decision node are shown. (b) Decision tree developed to separate HAM from healthy individuals. The *m/z* and intensity cutoff values are shown. (c) Decision tree developed by combining the two trees shown in (a) and (b). This tree was used as a test tree only and the reported value is the percent of samples that were correctly classified. The samples denoted with an asterisk are those samples that were misclassified.

have yet been identified that correspond to hematologic cancers. In this study, we have identified several of the component proteins/peptides that comprise the expression profile specific to ATL. Perhaps, as HTLV-I is a disease that infects circulating immune cells (predominantly CD4+), it is reasonable to suggest that a robust signature profile pattern may be more likely to be expressed in serum.

Both alpha-1-antitrypsin and haptoglobin-2 are acute-phase reactant proteins and as such play a role in a variety of cellular pathologies including diseases that involve inflammation response, hemolysis, and, in general, dramatic changes to blood homeostasis. Specifically, the serpin alpha-1-antitrypsin has been shown to be intimately involved in lymphocyte activation and metastasis.[18] Likewise, haptoglobin-2 is a serum protein which appears to be a natural suppressor of T-helper cell function as well as being a marker of T-cell activation.[19] In addition, haptoglobin levels and discrete protein modifications of haptoglobin are proving to be valuable markers for a variety of diseases of inflammation, infections, and neoplasia.[20–22] In our approach, we found specific stable fragments of these proteins to be useful in discriminating class, which suggest that these are disease-specific proteolytic events

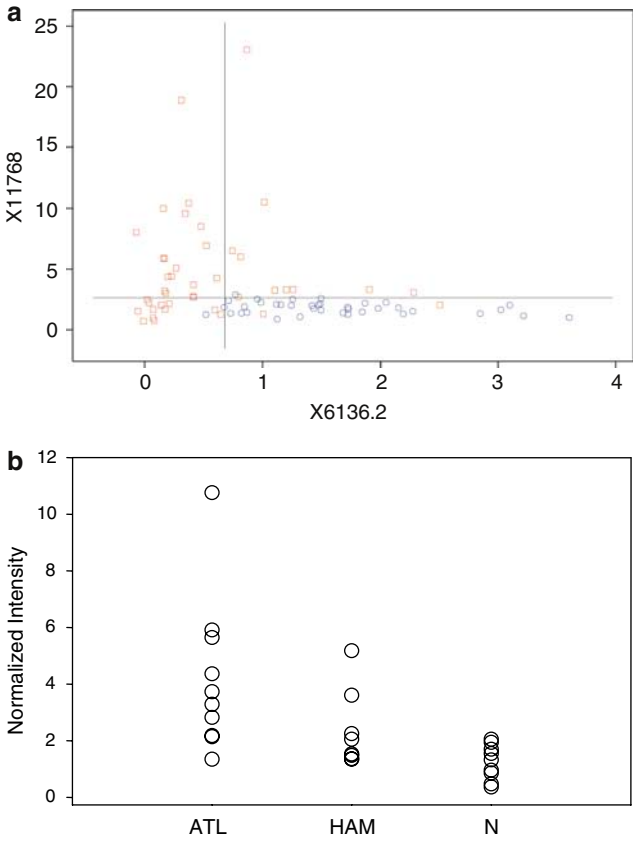

Figure 5 The 11 768 *m/z* peak is a robust discriminator. (a) A scatter plot representation of the ability of the 11 768 and 6136 *m/z* peaks to segregate between ATL and normal samples. The absolute normalized values for each peak found in ATL (red squares) and control (blue circles) were plotted on each axis. (b) The relative intensities for the 11 768 *m/z* peak were plotted for the blinded validation sample set. The sample set was derived from patients other than those used for the algorithm development.

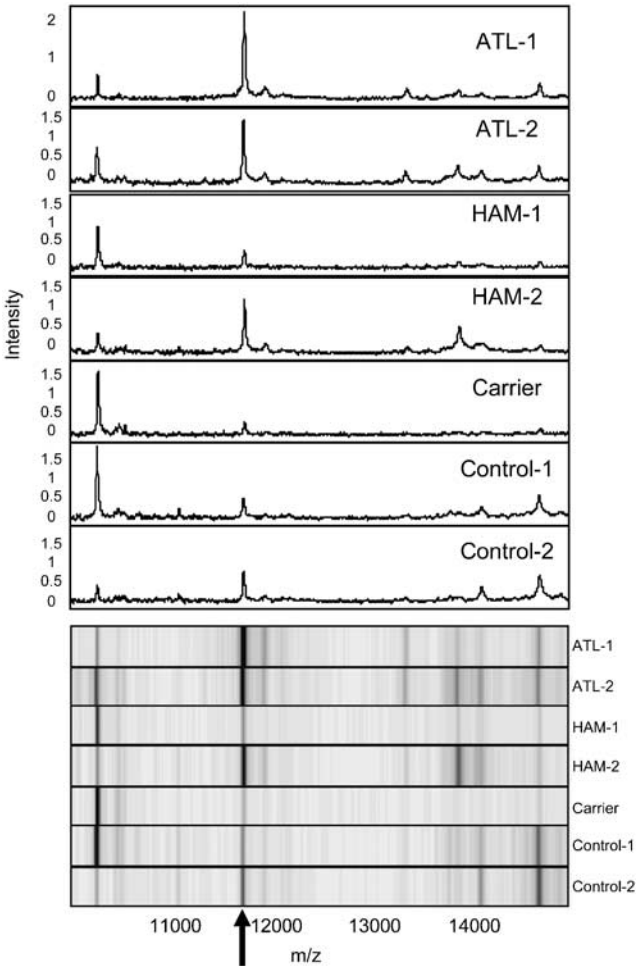

Figure 6 Expression profile and gel view of the region surrounding the *m/z* 11 768 peak. This peak (arrow) was present in all samples. Two samples from ATL, HAM, and control group and one from an asymptomatic carrier as indicated are shown. The upper panel is the spectral view and the lower panel is a gel view of the same data. The intensity values are relative to the normalized expression values.

that are amplified through the direct detection of enzymatic products.

Indeed, the 19.9 and 11.9 kDa fragments represent contiguous halves of the same protein, haptoglobin-2. Interestingly, a unique consensus site for a proline protease exists in haptoglobin-2, the cleavage of which would result in the production of two fragments of comparable mass. Our suggestion that an endogenous protease is the direct biomarker event is reminiscent of the observation of an increase in serum levels of the serine protease PSA as a marker for prostate cancer. Thus, the direct event is the increase in the expression of a specific protease and the secondary events would be the cleaved protein targets such as were identified in our study. It is interesting to speculate that the proteolytic fragments of the whole proteins may reveal more fine detail of hematopoietic changes that occur during leukemia. Specifically, the fragmentation pattern of a substrate protein for a leukemia specific protease may provide a more accurate map of disease development. The link between these proteins and cell homeostasis, immune response, and inflammation make them ideal candidates as putative biomarkers that signal early disease and potentially early signs of progression.

In summary, we have shown that protein expression profiling using SELDI is an accurate and sensitive approach to differentiating neurological from hematological disease following HTLV-I infection. We have achieved very high sensitivity and specificity in a blinded test of the algorithm, especially for ATL. The process appears to utilize the differential expression of fragments of relatively abundant sera proteins. These fragments are potentially the specific products of induced proteases, suggesting that the SELDI-assisted protein expression profiling approach monitors the integrity of classes of sera proteins that may be useful sentinels for disease development. We are actively investigating the identity of these proteases in the hopes of understanding disease development in ATL and to assist in improving HTLV-I disease diagnostics. In addition, we are undertaking an analysis of patient serum derived in a longitudinal study in order to evaluate the predictability of this marker for early disease detection and assessment of disease development. While not the focus of this report, protein expression profiling may be a powerful tool to determine if there are serum signature profiles and/or serum proteins that can predict disease outcomes in HTLV-I asymptomatic carriers. Specific to this clinical goal, we have recruited a large number of asymptomatic infected individuals, multiple sclerosis

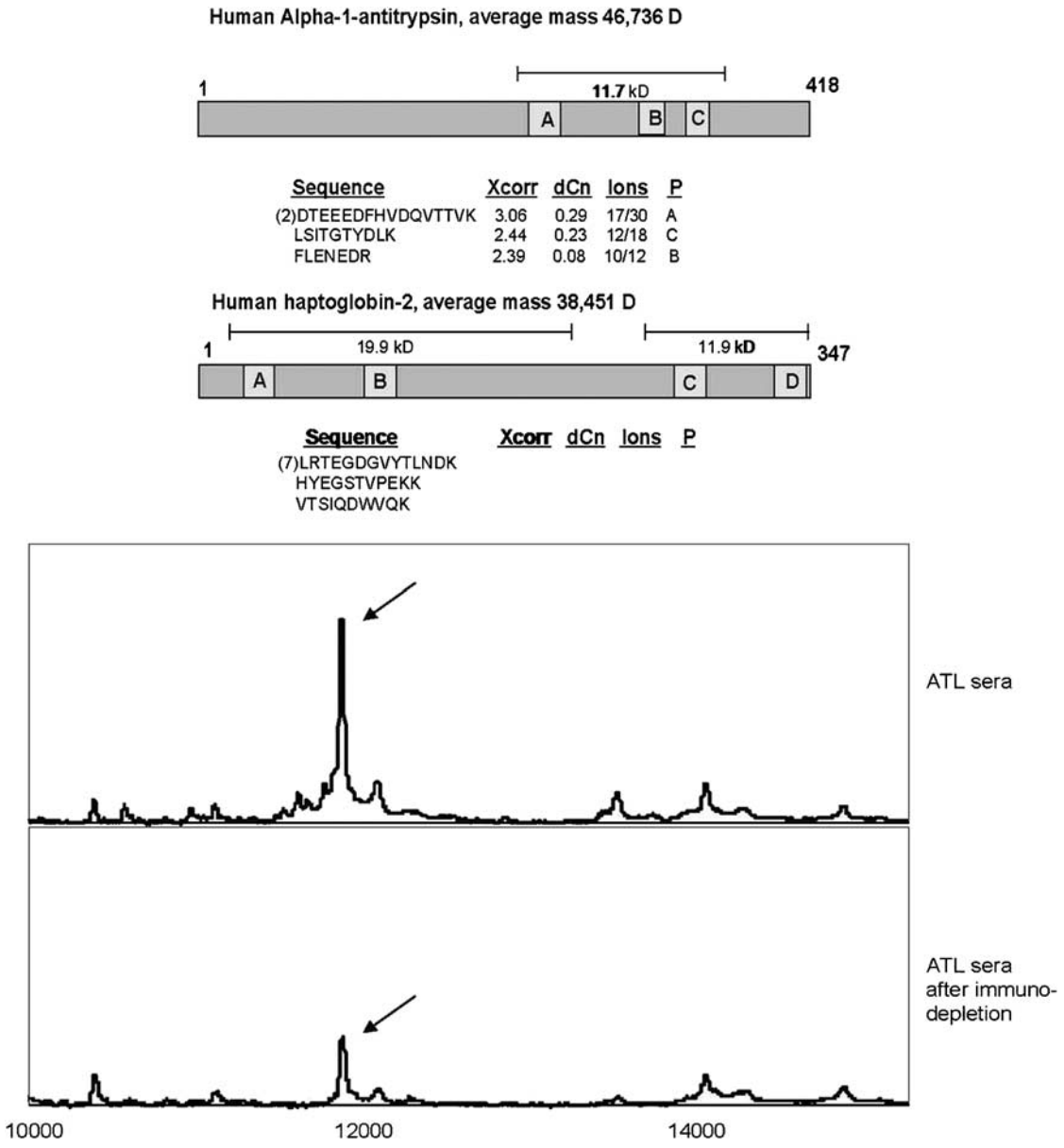

Figure 7 The resulting coverage of peptide identities from MS/MS and immunoassay results. (a) Three separate peptide 'hits' showing their position within the putative 11.7 kDa region of human alpha-1-antitrypsin. Each of the regions 'hit' within human haproglobin-2 were achieved with two separate peptides. For each peptide identified, the cross-correlation (Xcorr), delta-correlation (dCn), and the ion spread (Ions) are shown. (b) Immunodepletion of ATL sera with the α1-antitrypsin antibody. The upper trace is the untreated ATL sera and the bottom trace is the immunodepleted ATL sera. Note the decrease in the intensity of the peak at *m/z* 11 700.

patients, and other T-cell leukemia as comparative classes. We will employ our successive cohort design with blind validation test sets utilizing samples from multiple clinical sites in an effort to identify a robust biomarker of HTLV-I disease.

References

1 Hobbs SK, Shi G, Homer R, Harsh G, Atlas SW, Bednarski MD. Magnetic resonance image-guided proteomics of human glioblastoma multiforme. *J Magn Reson Imaging* 2003; **18**: 530–536.
2 Bangham CR. HTLV-1 infections. *J Clin Pathol* 2000; **53**: 581–586.
3 de The G, Bomford R. An HTLV-I vaccine: why, how, for whom? *AIDS Res Hum Retroviruses* 1993; **9**: 381–386.
4 Tajima K. The fourth nation-wide study of adult T-cell leukemia/lymphoma (ATL) in Japan: estimates of risk of ATL and its geographic and clinical features. *Int J Cancer* 1990; **45**: 237–243.
5 Pawson R, Richardson DS, Pagliuca A, Kelsey SM, Hoque S, Breuer J *et al*. Adult T-cell leukemia/lymphoma in London: clinical experience of 21 cases. *Leuk Lymphoma* 1998; **16**: 32–50.
6 Matutes E, Chan LC. Mixed-lineage leukemia revisited: acute lymphocytic leukemia with myeloperoxidase-positive blasts by electron microscopy. *Blood* 1991; **77**: 410–411.
7 Group LS. Major prognostic factors of patients with adult T-cell leukemia–lymphoma: a cooperative study. *Leuk Res* 1991; **15**: 81–90.
8 Mora CA, Osame M, Jacobson S. Human T lymphotrophic virus type-I associated myelopathy/tropical spastic paraparesis (HAM/TSP): therapeutic approach. *Curr Treat Options Infect Dis* 2003; **5**: 443–455.

9 Li J, Zhang Z, Rosenzweig J, Wang YY, Chan DW. Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. *Clin Chem* 2002; **48**: 1296–1304.
10 Adam BL, Qu Y, Davis JW, Ward MD, Clements MA, Cazares LH *et al*. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res* 2002; **62**: 3609–3614.
11 Wadsworth JT, Somers KD, Cazares LH, Malik G, Adam BL, Stack Jr BC *et al*. Serum protein profiles to identify head and neck cancer. *Clin Cancer Res* 2004; **10**: 1625–1632.
12 Vlahou A, Schellhammer PF, Mendrinos S, Patel K, Kondylis FI, Gong L *et al*. Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine. *Am J Pathol* 2001; **158**: 1491–1501.
13 Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM *et al*. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 2002; **359**: 572–577.
14 Qu Y, Adam BL, Yasui Y, Ward MD, Cazares LH, Schellhammer PF *et al*. Boosted decision tree analysis of surface-enhanced laser desorption/ionization mass spectral serum profiles discriminates prostate cancer from noncancer patients. *Clin Chem* 2002; **48**: 1835–1843.
15 Bertone P, Kluger Y, Lan N, Zheng D, Christendat D, Yee A *et al*. SPINE: an integrated tracking database and data mining approach for identifying feasible targets in high-throughput structural proteomics. *Nucleic Acids Res* 2001; **29**: 2884–2898.
16 Kosuda S, Ichihara K, Watanabe M, Kobayashi H, Kusano S. Decision-tree sensitivity analysis for cost-effectiveness of whole-body FDG PET in the management of patients with non-small-cell lung carcinoma in Japan. *Ann Nucl Med* 2002; **16**: 263–271.
17 Semmes OJ, Feng Z, Adam BL, Banez LL, Bigbee WL, Campos D *et al*. Evaluation of serum protein profiling by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry for the detection of prostate cancer: I. Assessment of platform reproducibility. *Clin Chem* 2005; **51**: 102–112.
18 Sun Z, Yang P. Role of imbalance between neutrophil elastase and alpha 1-antitrypsin in cancer development and progression. *Lancet Oncol* 2004; **5**: 182–190.
19 Arredouani M, Matthijs P, Van Hoeyveld E, Kasran A, Baumann H, Ceuppens JL *et al*. Haptoglobin directly affects T cells and suppresses T helper cell type 2 cytokine release. *Immunology* 2003; **108**: 144–151.
20 Langlois MR, Delanghe JR. Biological and clinical significance of haptoglobin polymorphism in humans. *Clin Chem* 1996; **42**: 1589–1600.
21 Turner GA. Haptoglobin A potential reporter molecule for glycosylation changes in disease. *Adv Exp Med Biol* 1995; **376**: 231–238.
22 Lee HB, Yoo OJ, Ham JS, Lee MH. Serum alpha 1-antitrypsin in patients with hepatocellular carcinoma. *Clin Chim Acta* 1992; **206**: 225–230.